132. The method of claim 131, wherein the analysis includes an analysis of hydrophobicity of the polypeptide, pka of one or more amino acids of the polypeptide, steric bulk of the polypeptide, or entropy of one or more amino acids of the polypeptide.

These amendments are made without prejudice and are not to be construed as abandonment of the previously claimed subject matter or agreement with any objection or rejection of record. In accordance with the requirements of 37 C.F.R. § 1.121, a marked up version showing the changes to the claims, is attached herewith as Appendix A. For the Examiner's convenience, a complete claim set of the currently pending claims is also submitted herewith as Appendix B.

REMARKS

STATUS OF THE CLAIMS.

Claims 93-132 are pending with entry of this amendment. The above amendments introduce no new matter, merely correcting antecedent basis issues and clarifying the claims. The amendments are not made for reasons of patentability and do not narrow the claims in any way. Indeed, the amendments broaden the scope of the claims by more clearly indicating that polypeptides as well as nucleic acids can be made according to the relevant methods and by changing "corresponding" language to "comprising" language. Support for the amendments to the claims is replete throughout the specification which clearly indicates throughout that the methods are applicable to polypeptides as well as nucleic acids.

New Claims 106-132 introduce no new matter. Support for the new claims is also replete throughout the claims and specification as filed. Example supporting passages include: page 7, line 14- page 9, line 4; page 17, line 14-line 23; page 38, line 13-line 18; page 47, line 14- page 64, line 2 (*e.g.*, page 60, line 17- page 64); and claims 93-105. Additional support is found throughout the specification which discusses, e.g., a wide variety of cross-over and bridging oligonucleotide formats.

Applicants also respectfully submit that new clams 106-132 are closely related to claims 93-105 and that no undue examination burden is believed to exist for consideration of the new claims with previously pending claims 93-105. The examination search for new claims 106-132 should be highly similar to the search already performed for claims 93-105 and the new claims should be grouped in the same class and subclass for search purposes. Because of the similarity of

the new claims to the previously pending claims, Applicants respectfully request that the new claims be entered and co-examined with previously pending claims 93-105.

Claims 94-95 were previously rejected for alleged lack of clarity with respect to the term "corresponding to." Claims 93-105 were also rejected for alleged obviousness over Jonsson et al. To the extent that the rejections are applied to the amended or new claims, Applicants traverse.

35 U.S.C. §112, SECOND PARAGRAPH.

Claims 94-95 were rejected under 35 U.S.C. §112, second paragraph, as allegedly being indefinite because of the use of the term "corresponding to." This has been amended (without narrowing of the claim) to the more common phrase "comprising," thereby removing any ambiguity from the claims. Accordingly, the rejection should be withdrawn.

THE CLAIMS ARE NOT OBVIOUS

Claims 93-105 were rejected under 35 U.S.C. §103(a) as allegedly obvious in light of Jonsson et al. (Nucleic Acids Research, 1993, 21(3):733-739). Applicants respectfully traverse.

As noted in the helpful Examiner's interview of August 8, 2001, Jonsson et al. is essentially unrelated to the claimed invention. That is, the Jonsson et al. reference relates to a way of designing and selecting sequences in silico that is fundamentally different from the methods which are now claimed.

In the claimed methods, parental sequences are provided and cross-over sites are designed based upon criteria of the practitioner. Nucleic acids are them made, e.g., using bridge oligonucleotides in various synthesis strategies.

This is simply not at all the approach used by Jonsson et al. That is, in the procedure of Jonsson et al., sequences are provided and each base residue in each sequence is assigned an indicator variable (A=1000, C=0100, G=0010 and T=0001; *see*, p. 733, column 2). All of the sequences are then placed in an indicator variable matrix, giving (in the relevant example) a 25x272 matrix (*see*, p. 735, column 1). PLS (partial least squares analysis) was then used to correlate a promoter efficiency variable (y) against the matrix (*see*, p. 735, column 2). Two PLS dimensions were found to account for 85% of the promoter efficiency variable (y) (*id.*). Particular residues (e.g., -12, 4-14, -8 to –10) were found to be the most significant. All of the systematic variance in the sequences were assessed via PLS (*see*, page 736, column 1 and Figure 3). Two of the sequences which came out of the PLS analysis were synthesized and tested for activity.

At no time, explicitly or inherently, were crossover sites between nucleic acids determined by Jonsson et al. Instead, all possible sites were considered as a matrix of information, against the promoter efficiency variable. Sequences of interest were suggested by PLS clustering that occurred for the sequences. This clustering was performed without consideration of cross-over information. In the end, nucleic acids were synthesized *without consideration of crossover information for any members of the matrix*. Because Jonsson et al. did not provide for consideration of cross-over events in their methods, the reference is essentially unrelated to the claims at issue.

It is also worth noting that no reason, articulated or otherwise, existed to modify the Jonsson methods, which operate independent of cross-over event consideration, to include consideration of cross-over events. That is, the Jonsson methods are asserted to be sufficient to provide for modeling of sequences *in silico*. There is no reason given in the Action, or apparent to Applicants, why one of skill would have completely changed the process proposed by Jonsson et al., absent the teaching in Applicants' disclosure.

Because Jonsson et al. does not teach the limitations of the claims at issue and because no reason is given to modify the procedures of Jonsson et al. to achieve the claimed invention, the rejection should be withdrawn.

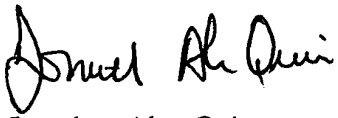
CONCLUSION

In view of the foregoing, Applicants believes all claims now pending in this application are in condition for allowance. The issuance of a formal Notice of Allowance at an early date is respectfully requested. **IN THE EVENT THAT ANY ISSUES REMAIN, APPLICANTS REQUEST A TELEPHONIC INTERVIEW PRIOR TO PREPARATION OF ANY ADDITIONAL ACTION BY THE OFFICE.**

If a telephone conference would expedite prosecution of this application, the Examiner is invited to telephone the undersigned at (510) 337-7871.

LAW OFFICES OF JONATHAN ALAN QUINE
P.O. BOX 458
Alameda, CA 94501
Tel: 510 337-7871
Fax: 510 337-7877

Respectfully submitted,



Jonathan Alan Quine
Reg. No: 41,261

APPENDIX A

"MARKED UP" CLAIMS ILLUSTRATING THE AMENDMENTS MADE TO THE CLAIMS OF 09/618,579 WITH ENTRY OF THIS AMENDMENT

93. **(AMENDED)** A method of producing recombinant nucleic acids or polypeptides, the method comprising:

providing two or more parental nucleic acid or polypeptide sequences;

selecting cross-over sites for recombination between the two or more parental nucleic acid or polypeptide sequences, thereby defining one or more recombinant nucleic acids or recombinant polypeptides that result from a cross-over between at least two of the two or more parental nucleic acids or polypeptides;

determining a recombinant sequence for at least one of the one or more recombinant nucleic acids or polypeptides;

selecting the at least one recombinant sequence in silico for one or more expected activity; and,

synthesizing the at least one recombinant sequence.

94 (AMENDED). The method of claim 93, further comprising [selecting] providing bridging oligonucleotides which [correspond to] comprise or encode the cross-over sites.

95 (AMENDED). The method of claim 94, wherein synthesizing the at least one recombinant sequence comprises providing fragments of the two or more parental nucleic acids and at least one of [corresponding] the bridge oligonucleotides, hybridizing the fragments and the bridge oligonucleotides and elongating the hybridized fragments with a polymerase or a ligase.

96. The method of claim 93, wherein the two or more parental sequences display low sequence similarity.

97 (AMENDED). The method of claim 93, wherein selecting the at least one recombinant sequence in silico comprises one or more of:

(i) performing an energy minimization analysis of the at least one recombinant sequence;

(ii) performing a stability analysis of the at least one recombinant sequence;

(iii) comparing an energy minimized model of the at least one recombinant sequence to an energy minimized model of one or more of the two or more parental nucleic acids or polypeptides;

(iv) performing protein threading on one or more [encoded protein] of the parental or recombinant polypeptides; and,

(v) selecting the cross-over sites for recombination between the two or more parental nucleic acid sequences or polypeptides to occur within regions of structural overlap, thereby determining the sequence of the at least one recombinant nucleic acid or polypeptide;

(vi) performing one or more of: PDA, a branch-and-terminate a combinatorial optimization analysis, a dead end elimination, a genetic or mean-field analysis, or analysis of protein folding by threading, of the at least one recombinant sequence;

(vii) performing PDA of at least one of the two or more parental sequences; or

(viii) comparing a PDA of the at least one recombinant sequence to a PDA of at least one of the two or more parental sequences.

98 (AMENDED). The method of claim 93, wherein the step of selecting cross-over sites for recombination between the two or more parental nucleic acid or polypeptide sequences and the step of selecting the at least one recombinant sequence in silico are performed simultaneously.

99 (AMENDED). A method of producing one or more recombinant nucleic acids or [encoded] polypeptides, the method comprising:

providing a plurality of first nucleic acid or first polypeptide sequences;

selecting cross-over sequences between the plurality of first nucleic acid or first polypeptide sequences by defining structural, statistical, or logical criteria for the cross-over sequences in silico; and,

artificially synthesizing a plurality of recombinant nucleic acids or polypeptides comprising or encoding the cross-over sequences.

100. The method of claim 99, wherein the first nucleic acid or polypeptide sequences comprise homologous sequences.

101. The method of claim 99, wherein the first nucleic acid or polypeptide sequences comprise non-homologous sequences.

102. The method of claim 99, wherein the first nucleic acid or polypeptide sequences comprise artificial sequences.

103. The method of claim 99, wherein the first nucleic acid or polypeptide sequences comprise sequences corresponding to naturally occurring nucleic acids or polypeptides.

104 (AMENDED). The method of claim 99, wherein defining the structural logical or statistical criteria comprises one or more of:

performing structural modeling of at least one of the first polypeptide sequences to define one or more region of structural interest in the at least one first polypeptide sequence and selecting one or more cross-over sequence to preserve or disrupt the region of structural interest;

defining a structural or sequence-based motif in at least one of the first polynucleotide or polypeptide sequences to define one or more conserved region in the at least one first polynucleotide or polypeptide sequence and selecting one or more cross-over sequence to preserve or disrupt the motif;

identifying one or more nucleotides or amino acids within at least one of the first polynucleotide or polypeptide sequences which shows activity or structural co-variance for one or more desired activities or structural features of the first polynucleotide or polypeptide sequence and selecting one or more cross-over sequence to preserve or disrupt the co-variance;

performing an energy minimization analysis of the first polynucleotide or polypeptide sequence and selecting one or more cross-over sequence to preserve or disrupt energy minimization of the first polynucleotide or polypeptide sequence;

performing a stability analysis of the first polynucleotide or polypeptide sequence and selecting one or more cross-over sequence to preserve or disrupt stability of the of the first polynucleotide or polypeptide sequence at least one recombinant sequence;

comparing an energy minimized model of the first polynucleotide or polypeptide sequence to an energy minimized model of one or more parental nucleic acid from which the first polynucleotide or polypeptide sequence was derived and selecting one or more cross-over sequence to preserve or alter energy minimization of the first polynucleotide or polypeptide sequence[.];

performing protein threading on one or more first polypeptide sequence and selecting the cross-over sequences to maintain or disrupt protein threading; and,

performing one or more of: PDA, a branch-and-terminate combinatorial optimization analysis, a dead end elimination, a genetic or mean-field analysis, or analysis of protein folding by threading, of the least one of the first polynucleotide or polypeptide sequence[;].

105 (AMENDED). The method of claim 99, wherein artificially synthesizing a plurality of recombinant nucleic acids comprising or encoding the cross-over sequences comprises synthesizing a plurality [or] <u>of</u> oligonucleotides, one or more of which encodes part or all of one or more of the cross-over sequences and incubating the plurality of oligonucleotides with a polymerase or a ligase, or both a polymerase and a ligase.

APPENDIX B

CLAIMS PENDING IN USSN 09/618,579 WITH ENTRY OF THIS AMENDMENT

93. **(AMENDED)** A method of producing recombinant nucleic acids or polypeptides, the method comprising:

providing two or more parental nucleic acid or polypeptide sequences;

selecting cross-over sites for recombination between the two or more parental nucleic acid or polypeptide sequences, thereby defining one or more recombinant nucleic acids or recombinant polypeptides that result from a cross-over between at least two of the two or more parental nucleic acids or polypeptides;

determining a recombinant sequence for at least one of the one or more recombinant nucleic acids or polypeptides;

selecting the at least one recombinant sequence in silico for one or more expected activity; and,

synthesizing the at least one recombinant sequence.

94 (AMENDED). The method of claim 93, further comprising providing bridging oligonucleotides which comprise or encode the cross-over sites.

95 (AMENDED). The method of claim 94, wherein synthesizing the at least one recombinant sequence comprises providing fragments of the two or more parental nucleic acids and at least one of [corresponding] the bridge oligonucleotides, hybridizing the fragments and the bridge oligonucleotides and elongating the hybridized fragments with a polymerase or a ligase.

96. The method of claim 93, wherein the two or more parental sequences display low sequence similarity.

97 (AMENDED). The method of claim 93, wherein selecting the at least one recombinant sequence in silico comprises one or more of:

(i) performing an energy minimization analysis of the at least one recombinant sequence;

(ii) performing a stability analysis of the at least one recombinant sequence;

(iii) comparing an energy minimized model of the at least one recombinant sequence to an energy minimized model of one or more of the two or more parental nucleic acids or polypeptides;

(iv) performing protein threading on one or more of the parental or recombinant polypeptides; and,

(v) selecting the cross-over sites for recombination between the two or more parental nucleic acid sequences or polypeptides to occur within regions of structural overlap, thereby determining the sequence of the at least one recombinant nucleic acid or polypeptide;

(vi) performing one or more of: PDA, a branch-and-terminate a combinatorial optimization analysis, a dead end elimination, a genetic or mean-field analysis, or analysis of protein folding by threading, of the at least one recombinant sequence;

(vii) performing PDA of at least one of the two or more parental sequences; or

(viii) comparing a PDA of the at least one recombinant sequence to a PDA of at least one of the two or more parental sequences.

98 (AMENDED). The method of claim 93, wherein the step of selecting cross-over sites for recombination between the two or more parental nucleic acid or polypeptide sequences and the step of selecting the at least one recombinant sequence in silico are performed simultaneously.

99 (AMENDED). A method of producing one or more recombinant nucleic acids or polypeptides, the method comprising:

providing a plurality of first nucleic acid or first polypeptide sequences;

selecting cross-over sequences between the plurality of first nucleic acid or first polypeptide sequences by defining structural, statistical, or logical criteria for the cross-over sequences in silico; and,

artificially synthesizing a plurality of recombinant nucleic acids or polypeptides comprising or encoding the cross-over sequences.

100. The method of claim 99, wherein the first nucleic acid or polypeptide sequences comprise homologous sequences.

101. The method of claim 99, wherein the first nucleic acid or polypeptide sequences comprise non-homologous sequences.

102. The method of claim 99, wherein the first nucleic acid or polypeptide sequences comprise artificial sequences.

103. The method of claim 99, wherein the first nucleic acid or polypeptide sequences comprise sequences corresponding to naturally occurring nucleic acids or polypeptides.

104 (AMENDED). The method of claim 99, wherein defining the structural logical or statistical criteria comprises one or more of:

performing structural modeling of at least one of the first polypeptide sequences to define one or more region of structural interest in the at least one first polypeptide sequence and selecting one or more cross-over sequence to preserve or disrupt the region of structural interest;

defining a structural or sequence-based motif in at least one of the first polynucleotide or polypeptide sequences to define one or more conserved region in the at least one first polynucleotide or polypeptide sequence and selecting one or more cross-over sequence to preserve or disrupt the motif;

identifying one or more nucleotides or amino acids within at least one of the first polynucleotide or polypeptide sequences which shows activity or structural co-variance for one or more desired activities or structural features of the first polynucleotide or polypeptide sequence and selecting one or more cross-over sequence to preserve or disrupt the co-variance;

performing an energy minimization analysis of the first polynucleotide or polypeptide sequence and selecting one or more cross-over sequence to preserve or disrupt energy minimization of the first polynucleotide or polypeptide sequence;

performing a stability analysis of the first polynucleotide or polypeptide sequence and selecting one or more cross-over sequence to preserve or disrupt stability of the of the first polynucleotide or polypeptide sequence at least one recombinant sequence;

comparing an energy minimized model of the first polynucleotide or polypeptide sequence to an energy minimized model of one or more parental nucleic acid from which the first polynucleotide or polypeptide sequence was derived and selecting one or more cross-over sequence to preserve or alter energy minimization of the first polynucleotide or polypeptide sequence;

performing protein threading on one or more first polypeptide sequence and selecting the cross-over sequences to maintain or disrupt protein threading; and,

performing one or more of: PDA, a branch-and-terminate combinatorial optimization analysis, a dead end elimination, a genetic or mean-field analysis, or analysis of protein folding by threading, of the least one of the first polynucleotide or polypeptide sequence.

105 (AMENDED). The method of claim 99, wherein artificially synthesizing a plurality of recombinant nucleic acids comprising or encoding the cross-over sequences comprises synthesizing a plurality of oligonucleotides, one or more of which encodes part or all of one or more of the cross-over sequences and incubating the plurality of oligonucleotides with a polymerase or a ligase, or both a polymerase and a ligase.

106 (NEW). A method of making a recombinant nucleic acid, the method comprising:

providing a plurality of parental character strings encoding a plurality of nucleic acids or polypeptides, which character strings, when aligned for maximum identity, comprise at least one region of heterology;

aligning the character strings;

selecting a plurality of cross-over sites in the character strings;

defining a set of character string subsequences, which set of subsequences comprises at least two subsequences from each of at least two of the plurality of parental character strings;

providing a set of oligonucleotides comprising or encoding the set of character string subsequences, which set of oligonucleotides comprises a plurality of bridging oligonucleotides which comprise or encode the plurality of cross-over sites;

annealing the set of oligonucleotides to each other; and,

elongating one or more members of the set of oligonucleotides with a polymerase, or ligating at least two members of the set of oligonucleotides with a ligase, or both elongating and ligating the set of oligonucleotides with a polymerase and a ligase, thereby producing one or more recombinant nucleic acid.

107 (NEW). The method of claim 106, further comprising expressing the recombinant nucleic acid to produce one or more recombinant polypeptide.

108 (NEW). The method of claim 106, wherein the two or more parental sequences display a sequence similarity of less than 50%.

109 (NEW). The method of claim 106, wherein the two or more parental sequences display a sequence similarity of more than 50%.

110 (NEW). The method of claim 106, further comprising determining one or more sequence for one or more putative recombinant nucleic acid resulting from in silico recombination of

the two or more parental sequences at the cross-over sites, and performing one or more in silico activity simulation for a polypeptide encoded by the putative recombinant nucleic acid.

111 (NEW). The method of claim 110, further comprising synthesizing the putative recombinant nucleic acid by providing fragments of the parental nucleic acids and at least one of the corresponding bridge oligonucleotides, hybridizing the fragments and the bridge oligonucleotides and elongating the hybridized fragments and/or bridge oligonucleotides with a polymerase and/or a ligase.

112 (NEW). The method of claim 106, wherein a parameter which is used in selecting a crossover site includes identification of one or more features of a nucleic acid corresponding to the character string, or of a polypeptide encoded by the nucleic acid corresponding to the character string, wherein the feature is selected from the group consisting of: structural stability, a 3-D energetic constraint, hydrophobicity, co-variation of residues, codon usage, motif distribution, one or more sequence motif, one or more active site, and one or more binding site.

113 (NEW). The method of claim 106, wherein a parameter which is used in selecting a crossover site includes an analysis of an amino acid composition of one or more polypeptide encoded by a nucleic acid that corresponds to one or more of the parental character strings.

114 (NEW). The method of claim 113, wherein the analysis includes an analysis of hydrophobicity of the polypeptide, pka of one or more amino acids of the polypeptide, steric bulk of the polypeptide, or entropy of one or more amino acids of the polypeptide.

115 (NEW). A method of producing recombinant nucleic acids, the method comprising:

providing two or more parental sequences;

selecting a plurality of cross-over sites for recombination between two or more of the parental sequences;

selecting a plurality of bridging oligonucleotides which correspond to the cross-over sites;

predicting at least one recombinant sequence defined by recombination between the parental sequences at the cross-over sites;

selecting the recombinant sequence in silico for one or more expected property of a recombinant nucleic acid that corresponds to the recombinant sequence, or one or more expected property of a polypeptide encoded by the recombinant nucleic acid; and,

synthesizing the recombinant nucleic acid.

116 (NEW). The method of claim 115, wherein synthesizing the recombinant nucleic acid comprises providing nucleic acid fragments which at least partly correspond to the two or more parental sequences and a plurality of bridge oligonucleotides, hybridizing the fragments and the bridge oligonucleotides and elongating the hybridized fragments with a polymerase or a ligase.

117 (NEW). The method of claim 115, wherein the parental sequences display sequence similarity of less than 50%.

118 (NEW). The method of claim 115, wherein selecting the recombinant sequence in silico comprises one or more of:

(i) performing an energy minimization analysis of a polypeptide encoded by a nucleic acid that corresponds to the recombinant sequence;

(ii) performing a stability analysis of a polypeptide encoded by a nucleic acid that corresponds to the recombinant sequence;

(iii) comparing an energy minimized model of a polypeptide encoded by a nucleic acid that corresponds to the recombinant sequence to an energy minimized model of a polypeptide encoded by a nucleic acid that corresponds to one or more of the parental sequences;

(iv) performing protein threading on a polypeptide encoded by a nucleic acid that corresponds to the parental or recombinant sequences;

(v) selecting the cross-over sites for recombination between the two or more parental sequences to occur within regions of structural overlap, thereby determining the sequence of the recombinant sequence;

(vi) performing one or more of: PDA, a branch-and-terminate a combinatorial optimization analysis, a dead end elimination, a genetic or mean-field analysis, or an analysis of polypeptide folding by threading, of the recombinant sequence or a polypeptide encoded by the recombinant sequence;

(vii) performing PDA of at least one of the parental sequences; or

(viii) comparing a PDA of the recombinant sequence to a PDA of at least one of the parental sequences.

119 (NEW). The method of claim 115, wherein the step of selecting cross-over sites for recombination between the parental sequences and the step of selecting the recombinant sequence in silico are performed simultaneously.

120 (NEW). The method of claim 115, wherein a parameter which is used in selecting at least one of the crossover sites includes identification of one or more features of one or more of the parental or recombinant sequences, or at least one polypeptide encoded by a nucleic acid corresponding to at least one of the parental or recombinant sequences, which feature is selected from the group consisting of: structural stability, a 3-D energetic constraint, hydrophobicity, co-variation of residues, codon usage, motif distribution, one or more sequence motif, one or more active site, and one or more binding site

121 (NEW). The method of claim 115, wherein a parameter which is used in selecting at least one crossover site includes the results of an analysis of an amino acid composition of a polypeptide encoded by a nucleic acid corresponding to the parental or recombinant sequence.

122 (NEW). The method of claim 121, wherein the analysis includes an analysis of hydrophobicity of the polypeptide, pka of one or more amino acids of the polypeptide, steric bulk of the polypeptide, or entropy of one or more amino acids of the polypeptide.

123 (NEW). A method of producing cross-over nucleic acids, the method comprising:

providing a plurality of first nucleic acid or first polypeptide sequences;

selecting a plurality of cross-over sequences between the plurality of first nucleic acid or first polypeptide sequences by defining structural, statistical, or logical criteria for the cross-over sequences in silico; and,

artificially synthesizing a plurality of recombinant nucleic acids comprising or encoding the plurality of cross-over sequences.

124 (NEW). The method of claim 123, wherein the first nucleic acid or polypeptide sequences comprise homologous sequences.

125 (NEW). The method of claim 123, wherein the first nucleic acid or polypeptide sequences comprise non-homologous sequences.

126 (NEW). The method of claim 123, wherein the first nucleic acid or polypeptide sequences comprise artificial sequences.

127 (NEW). The method of claim 123, wherein the first nucleic acid or polypeptide sequences comprise sequences corresponding to naturally occurring nucleic acids or polypeptides.

128 (NEW). The method of claim 123, wherein defining the structural logical or statistical criteria comprises one or more of:

performing structural modeling of at least one of the first polypeptide sequences to define one or more region of structural interest in the first polypeptide sequence and selecting one or more cross-over sequence to preserve or disrupt the region of structural interest;

defining a structural or sequence-based motif in at least one of the first polynucleotide or polypeptide sequences to define one or more conserved region in the first polynucleotide or polypeptide sequence and selecting one or more cross-over sequence to preserve or disrupt the motif;

identifying one or more nucleotides or amino acids within at least one of the first polynucleotide or polypeptide sequences which shows activity or structural co-variance for one or more desired activities or structural features of the first polynucleotide or polypeptide sequence, and selecting one or more cross-over sequence to preserve or disrupt the co-variance;

performing an energy minimization analysis of the first polynucleotide or polypeptide sequence and selecting one or more cross-over sequence to preserve or disrupt energy minimization of the first polynucleotide or polypeptide sequence;

performing a stability analysis of the first polynucleotide or polypeptide sequence and selecting one or more cross-over sequence to preserve or disrupt stability of the first polynucleotide or polypeptide sequence at least one recombinant sequence;

comparing an energy minimized model of the first polynucleotide or polypeptide sequence to an energy minimized model of one or more parental nucleic acid from which the first polynucleotide or polypeptide sequence was derived and selecting one or more cross-over sequence to preserve or alter energy minimization of the first polynucleotide or polypeptide sequence;

performing protein threading on one or more of the first polypeptide sequences and selecting the cross-over sequences to maintain or disrupt polypeptide threading; and,

performing one or more of: PDA, a branch-and-terminate combinatorial optimization analysis, a dead end elimination, a genetic or mean-field analysis, or analysis of polypeptide folding by threading, of the least one of the first polynucleotide or polypeptide sequence.

129 (NEW). The method of claim 123, further comprising incubating the plurality of cross-over nucleic acids with a polymerase or a ligase, or both a polymerase and a ligase.

130 (NEW). The method of claim 123, wherein a parameter which is used in selecting at least one of the crossover sites includes consideration of one or more features of at least one polynucleotide or polypeptide corresponding to at least one of the first polynucleotide or polypeptide sequences, which feature is selected from the group consisting of: structural stability, a

3-D energetic constraint, hydrophobicity, co-variation of residues, codon usage, motif distribution, one or more sequence motif, one or more active site, and one or more binding site.

131 (NEW). The method of claim 123, wherein a parameter which is used in selecting at least one crossover site includes the results of an analysis of an amino acid composition of a polypeptide corresponding to one or more of the first polypeptide sequences.

132 (NEW). The method of claim 131, wherein the analysis includes an analysis of hydrophobicity of the polypeptide, pka of one or more amino acids of the polypeptide, steric bulk of the polypeptide, or entropy of one or more amino acids of the polypeptide.

APPENDIX C

"MARKED UP" PARAGRAPHS ILLUSTRATING THE AMENDMENTS MADE TO THE SPECIFICATION OF 09/618,579 WITH ENTRY OF THIS AMENDMENT

Please delete the paragraph beginning at page 10, line 24 and ending at page 11, line 2 and substitute therefor the following new paragraph:

An introduction to genetic algorithms can be found in David E. Goldberg (1989) Genetic Algorithms in Search, Optimization and Machine Learning Addison-Wesley Pub Co; ISBN: 0201157675 and in Timothy Masters (1993) Practical Neural Network Recipes in C++ (Book&Disk edition) Academic Pr; ISBN: 0124790402. A variety of more recent references discuss the use of genetic algorithms used to solve a variety of difficult problems. *See*, e.g., **[http://]**garage.cse.msu.edu/papers/papers-index.html **(on the world wide web)** and the references cited therein; **[http://]**gaslab.cs.unr.edu/ **(on the world wide web)** and the references cited therein; **[http://www.]**aic.nrl.navy.mil/ **(on the world wide web)** and the references cited therein; **[http://www.]**cs.gmu.edu/research/gag/ **(on the world wide web)** and the references cited therein and **[http://www.]**cs.gmu.edu/research/gag/pubs.html **(on the world wide web)** and the references cited therein.

Please delete the paragraph beginning at page 18, line 26 and ending at page 19, line 17 and substitute therefor the following new paragraph:

One example algorithm that is suitable for determining percent sequence identity and sequence similarity is the BLAST algorithm, which is described in Altschul *et al., J. Mol. Biol.* 215:403-410 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (**[http://www.]**ncbi.nlm.nih.gov/ **on the world wide web**). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters

M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915).

Please delete the paragraph beginning at page 22, line 13 and ending at page 22, line 26 and substitute therefor the following new paragraph:

For example, oligonucleotides *e.g.*, for use in *in vitro* amplification/ gene reconstruction methods, for use as gene probes, or as shuffling targets (e.g., synthetic genes or gene segments) are typically synthesized chemically according to the solid phase phosphoramidite triester method described by Beaucage and Caruthers (1981), *Tetrahedron Letts.*, 22(20):1859-1862, *e.g.*, using an automated synthesizer, as described in Needham-VanDevanter *et al.* (1984) *Nucleic Acids Res.*, 12:6159-6168. Oligonucleotides can also be custom made and ordered from a variety of commercial sources known to persons of skill. There are many commercial providers of oligo synthesis services, and thus this is a broadly accessible technology. Any nucleic acid can be custom ordered from any of a variety of commercial sources, such as The Midland Certified Reagent Company (mcrc@oligos.com), The Great American Gene Company (**[http://www.]**genco.com**, on the world wide web**), ExpressGen Inc. (**[www.]**expressgen.com**, on the world wide web**), Operon Technologies Inc. (Alameda, CA) and many others. Similarly, peptides and antibodies can be custom ordered from any of a variety of sources, such as PeptidoGenic (pkim@ccnet.com), HTI Bio-products, inc. (**[http://www.]**htibio.com**, on the world wide web**), BMA Biomedicals Ltd (U.K.), Bio·Synthesis, Inc., and many others.

Please delete the paragraph beginning at page 44, line 28 and ending at page 45, line 8 and substitute therefor the following new paragraph:

If the assay conditions are then altered in only one parameter, different individuals from the library will be identified as the best performers. Because the screening conditions are very similar, most amino acids are conserved between the two sets of best performers. Comparisons of the sequences (e.g., in silico) of the best enzymes under the two different conditions identifies the sequence differences responsible for the differences in performance. Principal component analysis is a powerful tool to use for identifying sequences conferring a particular property. For example, Partek Incorporated (St. Peters, Missouri; **[www.]**partek.com**, on the world wide web**) provides software for pattern recognition (e.g., which provide Partek Pro 2000 Pattern Recognition Software) which can be applied to genetic algorithms for multivariate data analysis, interactive visualization, variable selection, neural & statistical modeling. Relationships can be analyzed, e.g., by Principal Components Analysis (PCA) mapped scatterplots and biplots, Multi-Dimensional Scaling (MDS) mapped scatterplots, Star plots, etc.

Please delete the paragraph beginning at page 47, line 5 and ending at page 47, line 13 and substitute therefor the following new paragraph:

For example, neural net approaches can be coupled to genetic algorithm-type programming. for example, NNUGA (Neural Network Using Genetic Algorithms) is an available program ([http://www.]cs.bgu.ac.il/~omri/NNUGA/**, on the world wide web**) which couples neural networks and genetic algorithms. An introduction to neural networks can be found, e.g., in Kevin Gurney (1999) An Introduction to Neural Networks, UCL Press, 1 Gunpowder Square, London EC4A 3DE, UK. and at **[http://www.]**shef.ac.uk/psychology/gurney/notes/index.html **(on the world wide web)** additional useful neural network references include those noted above in regard to genetic algorithms and, e.g., Christopher M. Bishop (1995) Neural Networks for Pattern Recognition Oxford Univ Press; ISBN: 0198538642; Brian D. Ripley, N. L. Hjort (Contributor) (1995) Pattern Recognition and Neural Networks Cambridge Univ Pr (Short); ISBN: 0521460867.

Please delete the paragraph beginning at page 50, line 23 and ending at page 51, line 20 and substitute therefor the following new paragraph:

With respect to modeling and structural analysis, a reductionist approach, in which protein positions are classified by their local environments, has aided development of appropriate energy expressions (reviewed in Street and Mayo (1999) "Computational Protein Design" Structure with Folding and Design 7(5):R105-R109). For examples of cycles of modeling and structural analysis, *See*, e.g., **[http://www.]**mayo.caltech.edu/ **(on the world wide web)**; Gordon and Mayo

(1999) "Branch-and-Terminate: A Combinatorial Optimization Algorithm for Protein Design" Structure with Folding and Design 7(9):1089-1098; Street and Mayo (1999) "Intrinsic ß-sheet Propensities Result from van der Waals Interactions Between Side Chains and the Local Backbone" Proc. Natl. Acad. Sci. USA, 96, 9074-9076; Gordon et al. (1999) "Energy Functions for Protein Design" Current Opinion in Structural Biology 9(4):509-513; Street and Mayo (1999) "Computational Protein Design" Structure with Folding and Design 7(5):R105-R109; Strop and Mayo (1999) "Rubredoxin Variant Folds Without Iron" J. Am. Chem. Soc. 121(11):2341-2345; Gordon and Mayo (1998) "Radical Performance Enhancements for Combinatorial Optimization Algorithms based on the Dead-End Elimination Theorem" J. Comp. Chem 19:1505-1514; Malakauskas and Mayo (1998) "Design, Structure, and Stability of a Hyperthermophilic Protein Variant" Nature Struct. Biol. 5:470; Street and Mayo (1998) "Pairwise Calculation of Protein Solvent-Accessible Surface Areas" Folding & Design 3: 253-258. Dahiyat and Mayo (1997) "De Novo Protein Design: Fully Automated Sequence Selection" Science 278:82-87; Dahiyat and Mayo (1997) "Probing the Role of Packing Specificity in Protein Design" Proc. Natl. Acad. Sci. USA 94:10172-10177; Haney et al. (1997) "Structural basis for thermostability and identification of potential active site residues for adenylate kinases from the archaeal genus *Methanococcus*" Proteins 28(1):117-30; and, Dahiyat et al. (1997) "Automated Design of the Surface Positions of Protein Helices" Prot. Sci. 6:1333-1337; Dahiyat et al. (1997) "De Novo Protein Design: Towards Fully Automated Sequence Selection" J. Mol. Biol. 273:789-796. Protein design programs can be used to build or modify proteins with any selected set of design criteria and these design criteria can be used as filters for any GO as noted herein, and/or in recursive cycles of design (by modeling or structural analysis), in silico GO or GA application, and/or physical recombination of nucleic acids of interest.

Please delete the paragraph beginning at page 55, line 6 and ending at page 55, line 18 and substitute therefor the following new paragraph:

Similarly, molecular dynamic simulations such as those above and, e.g., Ornstein et al. ([**http://www.**]emsl.pnl.gov:2080/homes/tms/bms.html **(on the world wide web)**; Curr Opin Struct Biol (1999) 9(4):509-13) provide for "rational" enzyme redesign by biomolecular modeling & simulation to find new enzymatic forms that would otherwise have a low probability of evolving biologically. For example, rational redesign of p450 cytochromes and alkane dehalogenase enzymes are a target of current rational design efforts. Any rationally designed protein (e.g., new p450 homologues or new alkaline dehydrogenase proteins) can be evolved by reverse translation and

shuffling against either other designed proteins or against related natural homologous enzymes. Details on p450s can be found in Ortiz de Montellano (ed.) 1995, Cytochrome P450 Structure and Mechanism and Biochemistry, Second Edition Plenum Press (New York and London). Furthermore, the dynamic simulations can be used as predictors of residues of interest and coupling or linkages between residues of interest and cross-over selection (or other GOs) can be performed to specifically maintain or eliminate such linkages in silico.

Please delete the paragraph beginning at page 64, line 28 and ending at page 65, line 6 and substitute therefor the following new paragraph:

HMM can be used in other ways as well. Instead of applying the generated profile to identify previously unidentified family members, the HMM profile can be used as a template to generate de novo family members (e.g., intermediate members of a cladistic tree of nucleic acids). For example, the program, HMMER is available (**[http://]**hmmer.wustl.edu/**, on the world wide web**). This program builds a HMM profile on a defined set of family members. A sub-program, HMMEMIT, reads the profile and constructs de novo sequences based on that. The original purpose of HMMEMIT is to generate positive controls for the search pattern, but the program can be adapted to the present invention by using the output as in silico generated progeny of a HMM profile defined shuffling. According to the present invention, oligonucleotides corresponding to these nucleic acids are generated for recombination, gene reconstruction and screening.

Please delete the paragraph beginning at page 73, line 3 and ending at page 73, line 23 and substitute therefor the following new paragraph:

Typically, PDA starts with a protein backbone structure and designs the amino acid sequence to modify the protein's properties, while maintaining it's three dimensional folding properties. Large numbers of sequences can be manipulated using PDA, allowing for the design of protein structures (sequences, subsequences, etc.). PDA is described in a number of publications, including, e.g., Malakauskas and Mayo (1998) "Design, Structure and Stability of a Hyperthermophilic Protein Variant" Nature Struc. Biol. 5:470; Dahiyat and Mayo (1997) "De Novo Protein Design: Fully Automated Sequence Selection" Science, 278, 82-87. DeGrado, (1997) "Proteins from Scratch" Science, 278:80-81; Dahiyat, Sarisky and Mayo (1997) "De Novo Protein Design: Towards Fully Automated Sequence Selection" J. Mol. Biol. 273:789-796; Dahiyat and Mayo (1997) "Probing the Role of Packing Specificity in Protein Design" Proc. Natl. Acad. Sci. USA, 94:10172-10177; Hellinga (1997) "Rational Protein Design – Combining Theory and Experiment" Proc. Natl. Acad.

Sci. USA, 94:10015-10017; Su and Mayo (1997)" Coupling Backbone Flexibility and Amino Acid Sequence Selection in Protein Design" Prot. Sci. 6:1701-1707; Dahiyat, Gordon and Mayo (1997) "Automated Design of the Surface Positions of Protein Helices" Prot. Sci., 6:1333-1337; Dahiyat and Mayo (1996) "Protein Design Automation" Prot. Sci., 5:895-903. Additional details regarding PDA are available, e.g., at [http://www.]xencor.com **(on the world wide web)**.

Please delete the paragraph beginning at page 80, line 27 and ending at page 81, line 2 and substitute therefor the following new paragraph:

4. Potential sequences for recombination can also be identified by computational methods other than by direct homology or structural information, e.g. by the PRINTS system (Attwood TK, Beck ME, Bleasby AJ, Degtyarenko K, Parry Smith DJ Nucleic Acids Res 1996 Jan 1;24(1):182-8), Shotgun (Pegg and Babbitt (1999) Bioinformatics 15(9):729-40) independently evolving sequence modules (may correspond to folding units as in MasterCatalog ([www.]eragen.com**, on the world wide web**) and other sequence comparison methods.

Please delete the paragraph beginning at page 81, line 9 and ending at page 81, line 17 and substitute therefor the following new paragraph:

Similarly, PRINTS (e.g., Atwood et al., *above*) is a compendium of protein motif fingerprints derived from the OWL composite sequence database. Fingerprints are groups of motifs within sequence alignments whose conserved nature allows them to be used as signatures of family membership. Fingerprints can provide improved diagnostic reliability over single motif methods by virtue of the mutual context provided by motif neighbors. The database is now accessible via the UCL Bioinformatics Server on **[http:@ www.]**biochem.ucl.ac.uk/bsm/dbbrowser/ **(on the world wide web)**. Atwood et al. describe the database, its compilation and interrogation software, and its Web interface. *See also*, Attwood et al. (1997) "Novel developments with the PRINTS protein fingerprint database" Nucleic Acids Res 25(1):212-7.

Please delete the paragraph beginning at page 88, line 22 and ending at page 89, line 3 and substitute therefor the following new paragraph:

One approach to screening diverse libraries is to use a massively parallel solid-phase procedure to screen cells expressing shuffled nucleic acids, e.g., which encode enzymes for enhanced activity. Massively parallel solid-phase screening apparatus using absorption, fluorescence, or FRET are available. *See*, e.g., United States Patent 5,914,245 to Bylina, et al. (1999); *see also*, **[http://www.]**kairos-scientific.com **(on the world wide web);** Youvan et al. (1999)

"Fluorescence Imaging Micro-Spectrophotometer (FIMS)" Biotechnology et alia<[www.]et-al.com (on the world wide web)> 1:1-16; Yang et al. (1998) "High Resolution Imaging Microscope (HIRIM)" Biotechnology et alia, <[www.]et-al.com (on the world wide web)> 4:1-20; and Youvan et al. (1999) "Calibration of Fluorescence Resonance Energy Transfer in Microscopy Using Genetically Engineered GFP Derivatives on Nickel Chelating Beads" posted at [www.]kairos-scientific.com (on the world wide web). Following screening by these techniques, sequences of interest are typically isolated, optionally sequenced and the sequences used as set forth herein to design new sequences for in silico or other shuffling methods.

Please delete the paragraph beginning at page 96, line 6 and ending at page 96, line 17 and substitute therefor the following new paragraph:

Generally the charts are schematics of arrangements for components, and of process decision tree structures. It is apparent that many modifications of this particular arrangement for DEGAGGS, e.g., as set forth herein, can be developed and practiced. Certain quality control modules and links, as well as most of the generic artificial neural network learning components are omitted for clarity, but will be apparent to one of skill. The charts are in a continuous arrangement, each connectable head-to tail. Additional material and implementation of individual GO modules, and many arrangements of GOs in working sequences and trees, as used in GAGGS, are available in various software packages. Suitable references describing exemplar existing software are found, e.g., at [http://www.]aic.nrl.navy.mil/galist/ (on the world wide web) and at http://www.]cs.purdue.edu/coast/archive/clife/FAQ/www/Q20_2.htm (on the world wide web). It will be apparent that many of the decision steps represented in Figs. 1-4 are performed most easily with the assistance of a computer, using one or more software program to facilitate selection/ decision processes.